Express Mail Label No:
EL844512243US
Date of Deposit: 4-17-01

ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER AND METHODS FOR DIAGNOSING AND TREATING GASTRIC CANCER

Related Applications

This application is a divisional of application serial number 08/896,164, filed July 17, 1997, now pending. This application claims the benefit under 35 U.S.C. §120 of application serial number 08/896,164.

Field of the Invention

This invention relates to the isolation of genes associated with gastric cancer, methods of diagnosing gastric cancer using these, as well as other genes which are known, as well as therapeutic approaches to treating such conditions.

Background of the Invention

It is fairly well established that many pathological conditions, such as infections, cancer, autoimmune disorders, etc., are characterized by the inappropriate expression of certain molecules. These molecules thus serve as "markers" for a particular pathological or abnormal condition. Apart from their use as diagnostic "targets", i.e., materials to be identified to diagnose these abnormal conditions, the molecules serve as reagents which can be used to generate diagnostic and/or therapeutic agents. A by no means limiting example of this is the use of cancer markers to produce antibodies specific to a particular marker. Yet another non-limiting example is the use of a peptide which complexes with an MHC molecule, to generate cytolytic T cells against abnormal cells.

Preparation of such materials, of course, presupposes a source of the reagents used to generate these. Purification from cells is one laborious, far from sure method of doing so. Another preferred method is the isolation of nucleic acid molecules which encode a particular marker, followed by the use of the isolated encoding molecule to express the desired molecule.

To date, two strategies have been employed for the detection of such antigens, in e.g., human tumors. These will be referred to as the genetic approach and the biochemical approach. The genetic approach is exemplified by, e.g., de Plaen et al., Proc. Natl. Acad. Sci.

USA 85: 2275 (1988), incorporated by reference. In this approach, several hundred pools of plasmids of a cDNA library obtained from a tumor are transfected into recipient cells, such as COS cells, or into antigen-negative variants of tumor cell lines. Transfectants are screened for the expression of tumor antigens via their ability to provoke reactions by anti-tumor cytolytic T cell clones. The biochemical approach, exemplified by, e.g., Mandelboim, et al., Nature 369: 69 (1994) incorporated by reference, is based on acidic elution of peptides which have bound to MHC-class I molecules of tumor cells, followed by reversed-phase high performance liquid chromography (HPLC). Antigenic peptides are identified after they bind to empty MHC-class I molecules of mutant cell lines, defective in antigen processing, and induce specific reactions with cytolytic T-lymphocytes ("CTLs"). These reactions include induction of CTL proliferation, TNF release, and lysis of target cells, measurable in an MTT assay, or a ^{51}Cr release assay.

These two approaches to the molecular definition of antigens have the following disadvantages: first, they are enormously cumbersome, time-consuming and expensive; second, they depend on the establishment of CTLs with predefined specificity; and third, their relevance in vivo for the course of the pathology of disease in question has not been proven, as the respective CTLs can be obtained not only from patients with the respective disease, but also from healthy individuals, depending on their T cell repertoire.

The problems inherent to the two known approaches for the identification and molecular definition of antigens is best demonstrated by the fact that both methods have, so far, succeeded in defining only very few new antigens in human tumors. See, e.g., van der Bruggen et al., Science 254: 1643-1647 (1991); Brichard et al., J. Exp. Med. 178: 489-495 (1993); Coulie, et al., J. Exp. Med. 180: 35-42 (1994); Kawakami, et al., Proc. Natl. Acad. Sci. USA 91: 3515-3519 (1994).

Further, the methodologies described rely on the availability of established, permanent cell lines of the cancer type under consideration. It is very difficult to establish cell lines from certain cancer types, as is shown by, e.g., Oettgen, et al., Immunol. Allerg. Clin. North. Am. 10: 607-637 (1990). It is also known that some epithelial cell type cancers are poorly susceptible to CTLs in vitro, precluding routine analysis. These problems have stimulated the art to develop additional methodologies for identifying cancer associated antigens.

One key methodology is described by Sahin, et al., Proc. Natl. Acad. Sci. USA 92: 11810-11913 (1995), incorporated by reference. Also, see U.S. Patent Applications Serial

No. 08/580,980, and Application Serial No. 08/479,328, filed on June 7, 1995 and January 3, 1996, respectively. All three of these references are incorporated by reference. To summarize, the method involves the expression of cDNA libraries in a prokaryotic host. (The libraries are secured from a tumor sample). The expressed libraries are then immunoscreened with absorbed and diluted sera, in order to detect those antigens which elicit high titer humoral responses. This methodology is known as the SEREX method ("Serological identification of antigens by Recombinant Expression Cloning"). The methodology has been employed to confirm expression of previously identified tumor associated antigens, as well as to detect new ones. See the above referenced patent applications and Sahin, et al., supra, as well as Crew, et al., EMBO J 144: 2333-2340 (1995).

The SEREX methodology has now been applied to stomach cancer samples. Several nucleic acid molecules have been newly isolated and sequenced, and are now associated with stomach cancer. Further, a pattern of expression involving these, as well as previously isolated genes has been found to be associated with stomach cancer. These results are the subject of this application, which is elaborated upon in the disclosure which follows.

Detailed Description of Preferred Embodiments

Example 1

A stomach cancer cDNA library was established, using standard techniques, then the library was screened, using the SEREX methodology described supra, and set forth by Sahin et al., Proc. Natl. Acad. Sci. USA 92: 11810 (1995), and by Chen et al., Proc. Natl. Acad. Sci. USA 94: 1914 (1997), incorporated by reference in their entirety.

To be specific, total RNA was isolated by homogenizing tumor samples in 4M guanidium thiocyanate/0.5% sodium N-lauryl sarcosine/ and 25 mM EDTA followed by centrifugation in 5.7 M CsCl/25 mM sodium acetate/10 uM EDTA at 320,000 rpm. Total mRNA was removed by passing the sample over an oligo-dT cellulose column. The cDNA libraries were then constructed by taking 5 µg of mRNA, using standard methodologies to reverse transcribe the material.

Libraries were prepared from three different patients, referred to as "SM", "CK" and "SS" respectively. A total of $2.5x10^6$, $1.1x10^6$, and $1.7x10^6$ cDNA clones were obtained from the individuals.

The cDNA was used to construct a lambda phage library, and 500 phages were plated

onto XL1-Blue MRF E. coli, and incubated for eight hours at 37°C. A nitrocellulose membrane was then placed on the plate, followed by overnight incubation. The membrane was then washed, four times, without TBS which contained 0.05% Tween, and was then immersed in TBS containing 5% non-fat dried milk. After one hour, the membrane was incubated with conjugates of peroxidase-goat anti human IgG specific for Fc portions of human antibody (1:2000, diluted in TBS with 1% BSA. The incubation was carried out for one hour, at room temperature, and the membrane was then washed three times with TBS. Those clones which produced antibodies were visualized with 0.06%, 3,3'diamino benzidine tetrachloride, and 0.015% H_2O_2, in 50 mM Tris (pH 7.5). Any clones which produced immunoglobulin were marked, and then the membrane was washed, two further times, with TBS that contained 0.05% Tween, and then twice with "neat" TBS.

The membranes were then incubated in 1:100 diluted patient serum, overnight, at 4°C. The patient serum had been pretreated. Specifically, 5 ml samples were diluted to 10 ml with TBS containing 1% bovine serum albumin, and 0.02% Na_3N. The serum had been treated to remove antibodies to bacteriophage, by passing it through a 5 ml Sepharose column, to which a lysate of E. coli Y1090 had been attached, followed by passage over a second column which had E. coli lysate and lysate of E. coli infected with lambda bacteriophage. The screening was carried out five times. The samples were then diluted to 50 ml, and kept at -80°C, until used as described herein.

Following the overnight incubation with the membrane, the membrane was washed twice with TBS/0.05% Tween 20, and then once with TBS. A further incubation was carried out, using the protocols discussed supra, for the POD labelled antibodies.

In the case of library SM, 55 positive clones were obtained. Library CK yielded 56 positive clones after the first screening, and after two additional screenings, this number was reduced to 26.

The 55 positive clones were then sequenced, using standard techniques. Following comparison of the sequences to information available in data banks, a total of 36 clones were resolved into known and unknown genes. In the table that follows, the "+" and "-" signs are essentially used to compare signals to each other. All were positive. Table 1, which follows, summarizes some of this work. Specifically, with reference to the first page of the table, previously identified human proteins and the nucleotide sequences, set forth in SEQ ID NOS: 1 to 39 are known. The four molecules which follow (gelsolin, zinc finger protein family,

variant zinc finger motif protein goliath and homeodomain proteins, have not been identified in humans previously, although there are related molecules found in other species. SEQ ID NOS: 40 and 48 set forth these nucleotides sequences. Finally, with reference to the table the last four moieties, i.e., prepro-α collagen, heterogeneous ribonucleoprotein D, nucleosome assembly protein 2, and NY-ESO-2/Ulsn NRP/V1 small nuclear ribonucleoprotein, are also known. Nucleotide sequences are set forth at SEQ ID NOS: 49 to 55. The nucleic acid molecules having the nucleotide sequences set forth at SEQ ID NOS: 56 to 83 represent molecules for which no related sequences were found. SEQ ID NO: 84 combines the sequences of SEQ ID NOS: 40 to 43, inclusive. SEQ ID NO: 85 combines SEQ ID NOS: 56-69, SEQ ID NO: 86 combines SEQ ID NOS: 70, 72 and 75, while SEQ ID NO: 87 combines SEQ ID NOS: 71, 73, 74 and 76.

Table 1

	Clone	SEQ ID NO	Reactivity	Insert size
RPB-Jκ	S366	1, 2	++	1.75
Human H-2K	S224	3	++	1.56
Binding Factor 2	S541	4	++	1.98
(7 clones)	S584	5	++	1.86
	S194	6	++ - +	2.43
	S344	7, 8	+	1.69
	S699	9	+	1.85
Telomeric Repeat Binding	S385	10, 11	++	0.96
Protein (2 clones)	S4510	12	++	0.96
Protein Kinase B/AKT/	S621	13	++	2.00
Serine Threonine Protein	S641	14	+	1.80
Kinase rac alpha (2 clones)				
SRY Interacting Protein-1/	S72	15, 16	++	1.70
Tyrosine Kinase Activator	S321	17, 18	++	1.76
Protein (2 clones)				
Sterol Carrier Protein-X/	S571	19	++	2.20
Sterol Carrier Protein-2	S613	20	++	2.25
(3 clones)	S62	21, 22	+	2.91
Archain/Coat Protein	S537	23, 24	++	3.42
Delta-Cop (Bovine)	S632	25	++	0.90
(3 clones)	S591	26	++ - +	3.50

HEM-1 (2 clones)	S551 S634	27, 28 29	++ ++	1.95 3.46
Id-1 Helix-Loop-Helix Protein (1 clone)	S271	30, 31	+++	1.21
E2A Helix-Loop-Helix Transcription Factor (1 clone)	S231	32, 33	++	1.66
Follistatin-Related Protein (1 clone)	S706	34	+	3.60
Translation Initiation Factor eIF-4γ (1 clone)	S274	35, 36	+	3.82
M Phase Phosphoprotein 1 (1 clone)	S204	37, 38	+	4.71
Lysyl tRNA Synthetase (1 clone)	S691	39	+	1.80
Gelsolin Family (2 clones)	S564 S26	40, 41 42, 43	++ +	2.04 2.57
Zinc Finger Protein Family (1 clone)	S343	44, 45	++ - +	0.90
A Variant Zinc-Finger-Motif Protein Goliath	S622	46	++ - +	1.70
Homeodomain Protein Family (1 clone)	S4611	47, 48	+	3.20
Prepro-α1 (1) Collagen (1 clone)	S563	49	++	3.28
Heterogeneous Ribonucleoprotein D (3 clones)	S292 S222 S232	50 51 52	+ + +	1.25 2.49 1.60
Nucleosome Assembly Protein 2 (1 clone)	S524	53, 54	+	1.52
NY-ESO-2/U1sn NRP U1 Small Nuclear	S623	55	+	0.9

Ribonucleoprotein
(1 clone)

The foregoing examples demonstrate several features of the invention. These include diagnostic methods for determining presence of transformed cells, such as gastric cancer cells, in a sample. The sample may contain whole cells or it may be, e.g., a body fluid sample, or an effusion, etc., where the sample may contain cells, but generally will contain shed antigen. The experiments indicate that there is a family of proteins, expression of which is associated with gastric cancer. Hence, the invention involves, inter alia, detecting at least two of the proteins set out in Table 1 wherein, presence of these is indicative of a pathology, such as gastric cancer or other type of related condition. Exemplary of the type of diagnostic assays which can be carried out are immunoassays, amplification assays (e.g., PCR), or, what will be referred to herein as a "display array". "Display array" as used herein refers to a depiction of the protein profile of a given sample. Exemplary of such displays are 2-dimensional electrophoresis, banding patterns such as SDS-gels, and so forth. Thus, one aspect of the invention involves diagnosing gastric cancer or a related condition by determining protein display of a sample, wherein a determination of at least two of the proteins, or expression of their genes, as set forth in Table 1, is indicative of gastric cancer or a related condition. There are many ways to carry out these assays. For example, as indicated herein, antibodies to the proteins were found in patient samples. One can assay for these antibodies using, e.g., the methodology described herein, or by using a purified protein or proteins or antigenic fragment thereof, and so forth. One can also assay for the protein itself, using antibodies, which may be isolated from samples, or generated using the protein and standard techniques. This antibodies can then be labelled, if desired, and used in standard immunoassays. These antibodies or oligonucleotide probes/primers may also be used to examine biopsied tissue samples, e.g., to diagnose precancerous conditions, early stage cancers, and so forth.

Similarly, any and all nucleic acid hybridization systems can be used, including amplification assays, such as PCR, basic probe hybridization assays, and so forth. The antibodies, such as polyclonal antibodies, monoclonal antibodies, the hybridomas which produce them, recombinantly produced antibodies, binding fragments of these, hybridization kits, DNA probes, and so forth, are all additional features of the invention.

Any of these assays can also be used in progression/regression studies. One can monitor the course of an abnormality such as gastric cancer which involve expression of any one of the proteins, the expression of which is governed by nucleic acid molecules which comprise SEQ ID NOS: 40-48 and 56-87, simply by monitoring levels of the protein, its expression, and so forth using any or all of the methods set forth supra.

As has been indicated supra, the isolated nucleic acid molecules which comprise the nucleotide sequences set forth in SEQ ID NOS: 40-48 and 56-87 are new, in that they have never been isolated before. These nucleic acid molecules may be used as a source to generate gastric cancer specific proteins and peptides derived therefrom, and oligonucleotide probes which can themselves be used to detect expression of these genes. Hence, a further aspect of the invention is an isolated nucleic acid molecule which comprises any of the nucleotide sequences set forth in SEQ ID NOS: 40-48 and 56-87, or molecules whose complements hybridize to one or more of these nucleotide sequences, under stringent conditions, expression vectors comprising these molecules, operatively linked to promoters, cell lines and strains transformed or transfected with these, and so forth. "Stringent conditions", is used herein, refers to condition such as those specified in U.S. Patent No. 5,342,774, i.e., 18 hours of hybridization at 65°C, followed by four one hour washes at 2xSSC, 0.1% SDS, and a final wash at 0.2xSSC, more preferably 0.1xSSC, 0.1% SDS for 30 minutes, as well as alternate conditions which afford the same level of stringency, and more stringent conditions.

It should be clear that these methodologies may also be used to track the efficacy of a therapeutic regime. Essentially, one can take a baseline value for the protein or proteins being tested, using any of the assays discussed supra, administer a given therapeutic, and then monitor levels of the protein or proteins thereafter, observing changes in protein levels as indicia of the efficacy of the regime.

For those proteins which have been identified previously, i.e., those in Table 1, similar assays can be carried out, but for two or more of the proteins or nucleic acid molecules encoding them. Specifically, tilomeric repeat binding protein, or protein kinase B/AKT/serine threonine protein kinase rac alpha are preferably determined, either in the form of the protein per se, or nucleic acid molecules encoding them (e.g., SEQ ID NOS: 10-14), together, or in combination with one or more of human gelsolin (SEQ ID NOS: 40-43 and 84), one or more of SEQ ID NOS: 56-76, 85, 86 or 87), or the proteins encoded by these sequences, and so forth.

The identification of the proteins set forth herein as being implicated in pathological conditions such as gastric cancer also suggests a number of therapeutic approaches to such conditions. The experiments set forth supra establish that antibodies are produced in response to expression of these proteins, suggesting their use as a vaccine. Hence, a further embodiment of the invention is the treatment of conditions which are characterized by expression of one or more of the subject proteins, via immunotherapeutic approaches. One of these approaches is the administration of an amount of one or more these proteins, or an immunogenic peptide derived therefrom in an amount sufficient to provoke or augment an immune response. The proteins or peptides may be combined with one or more of the known immune adjuvants, such as saponins, GM-CSF, interleukins, and so forth. If the peptides are too small to generate a sufficient antibody response, they can be coupled to the well known conjugates used to stimulate responses.

Similarly, the immunotherapeutic approaches include administering an amount of inhibiting antibodies sufficient to inhibit the protein or proteins. These antibodies may be, e.g., antibodies produced via any of the standard approaches elaborated upon supra.

T cell responses may also be elicited by using peptides derived from the proteins which then complex, non-covalently, with MHC molecules, thereby stimulating proliferation of cytolytic T cells against any such complexes in the subject. It is to be noted that the T cells may also be elicited in vitro, and then reperfused into the subject being treated.

Note that the generation of T cells and/or antibodies can also be accomplished by administering cells, preferably treated to be rendered non-proliferative, which present relevant T cell or B cell epitopes for response.

The therapeutic approaches may also include gene therapies, wherein an antisense molecule, preferably from 10 to 100 nucleotides in length, is administered to the subject either "neat" or in a carrier, such as a liposome, to facilitate incorporation into a cell, followed by inhibition of expression of the protein. Such antisense sequences may also be incorporated into appropriate vaccines, such as in viral vectors (e.g., Vaccinia), bacterial constructs, such as variants of the well known BCG vaccine, and so forth.

An additional DNA based therapeutic approach is the use of a vector which comprises one or more nucleotide sequences, preferably a plurality of these, each of which encodes an immunoreactive peptide derived from the expressed proteins. One can combine these peptides expressing sequences in all possible variations, such as one from each protein,

several from one or more protein and one from each of the additional proteins, a plurality from some and none from others, and so forth.

Other features of the invention will be clear to the skilled artisan, and need not be repeated here.

The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, it being recognized that various modifications are possible within the scope of the invention.